Broadcast Channels

Sunil Srinivasa

1 Introduction

A broadcast channel (BC) is a communication channel that has one sender and two or more receivers. The BC with two receivers is illustrated in Figure 1. The study of BCs was initialized by Cover [1]. Simple examples of real-world BCs include TV or radio stations and lecturing in a classroom. The basic problem that we are looking at is to find the set of simultaneously achievable rates (R_1, R_2) for communication in such a channel. In this report, we discuss some of the progress towards finding the capacity region for the BC. Rather surprisingly, the capacity region for a general BC is still unknown and remains an open problem !

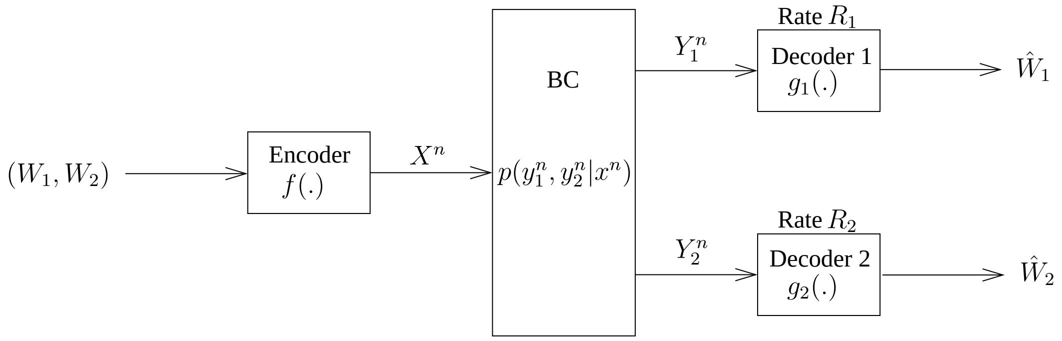

Figure 1: BC with two receivers. (W_1, W_2) is the transmitted message and $\hat{W}_1$ and $\hat{W}_2$ are the decoder estimates.

2 Definitions for a Broadcast Channel

The formal definitions for the BC with two receivers are as follows.

Definition A BC consists of an input alphabet $\mathcal{X}$ and two output alphabets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ and a probability transition function $p(y_1^n, y_2^n|x^n)$. The BC is said to be *memoryless* if $p(y_1^n, y_2^n|x^n) = \prod_{i=1}^n p(y_{1i}, y_{2i}|x_i)$.

Definition A $(\lceil 2^{nR_1} \rceil, \lceil 2^{nR_2} \rceil, n)$ code for a BC with independent information consists of an encoder,

$$f : (\{1, 2, \ldots, \lceil 2^{nR_1} \rceil\} \times \{1, 2, \ldots, \lceil 2^{nR_2} \rceil\}) \to \mathcal{X}^n, \tag{1}$$

and two decoders

$$g_i : \mathcal{Y}_i^n \to \{1, 2, \ldots, \lceil 2^{nR_i} \rceil\}, \quad i = 1, 2. \tag{2}$$

Definition The average probability of error is defined as the probability that either of the decoded messages is not equal to the corresponding transmitted message, i.e.,

$$P_e^{(n)} = \Pr \Big[g_1(Y_1^n) \neq W_1 \text{ or } g_2(Y_2^n) \neq W_2 \Big], \tag{3}$$

where (W_1, W_2) are assumed to be uniformly distributed over $\lceil 2^{nR_1} \rceil \times \lceil 2^{nR_2} \rceil$.

Definition A rate pair (R_1, R_2) is said to be *achievable* for the BC if there exists a sequence of $(\lceil 2^{nR_1} \rceil, \lceil 2^{nR_2} \rceil, n)$ codes with $P_e^{(n)} \to 0$ as $n \to \infty$[1].

Definition The *capacity region* of the BC is the closure of the set of achievable rates.

3 Degraded Broadcast Channel

It is often the case in practice that one receiver may experience a better channel than the other, in a certain sense. Such a channel is termed a *degraded* BC (Figure 2). This is one of the few classes of channels for which the capacity region is established.

Definition A memoryless BC is said to be *physically degraded* if $p(y_1, y_2|x) = p(y_1|x)p(y_2|y_1)$ i.e., $X \leftrightarrow Y_1 \leftrightarrow Y_2$.

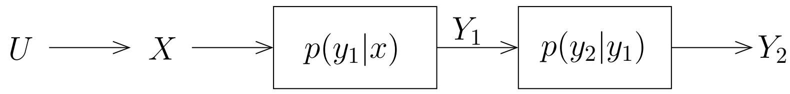

Figure 2: Degraded BC with auxiliary input U.

3.1 Capacity region for the Degraded Broadcast Channel

Theorem 3.1. *The capacity region for sending independent information over the degraded BC $X \leftrightarrow Y_1 \leftrightarrow Y_2$ is the closure of all (R_1, R_2) satisfying*

$$\begin{aligned} R_1 &< I(X; Y_1|U), \\ R_2 &< I(U; Y_2) \end{aligned} \tag{4}$$

for some joint distribution $p(u)p(x|u)p(y_1, y_2|x)$, where the auxiliary random variable U has cardinality bounded by $|\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}_1|, |\mathcal{Y}_2|\}$.

Proof. The proof hinges on the idea of superposition coding for the BC [2]. The main idea is to superimpose the message intended for the better receiver on the poorer receiver. The auxiliary random variable U serves as a "cloud center" distinguishable to both receivers Y_1 and Y_2. Each cloud consists of $\lceil 2^{nR_1} \rceil$ "satellite" codewords distinguishable only by receiver Y_1.

Outline of achievability:

- Fix $p(u)$ and $p(x|u)$.
- *Random codebook generation:* Generate $\lceil 2^{nR_2} \rceil$ independent codewords of length n, $u^n(w_2), w_2 \in \{1, 2, \ldots, \lceil 2^{nR_2} \rceil\}$, according to $\prod_{i=1}^n p(u_i)$. For each codeword $u^n(w_2)$, generate $\lceil 2^{nR_1} \rceil$ independent codewords $x^n(w_1, w_2)$ according to $\prod_{i=1}^n p(x_i|u_i(w_2))$. Thus, $u^n(i)$ plays the role of the cloud center understandable to both Y_1 and Y_2, while $x^n(i, j)$ is the j^{th} satellite codeword in the i^{th} cloud.
- *Encoding:* To send the pair (W_1, W_2), send the corresponding codeword $x^n(W_1, W_2)$.

[1]Under certain circumstances, the sender might be interested in transmitting common information to both receivers. A $(\lceil 2^{nR_0} \rceil, \lceil 2^{nR_1} \rceil, \lceil 2^{nR_2} \rceil, n)$ code for a BC with independent information consists of an encoder $f : (\{1, 2, \ldots, \lceil 2^{nR_0} \rceil\} \times \{1, 2, \ldots, \lceil 2^{nR_1} \rceil\} \times \{1, 2, \ldots, \lceil 2^{nR_2} \rceil\}) \to \mathcal{X}^n$ and two decoders $g_i : \mathcal{Y}_i^n \to \{1, 2, \ldots, \lceil 2^{nR_0} \rceil\} \times \{1, 2, \ldots, \lceil 2^{nR_i} \rceil\}, i = 1, 2$, where R_0 is the transmission rate for the common information. The definitions for the average probability of error and the rate triple are similar to those defined above.

- *Decoding*: The decoders perform joint typicality decoding based on the received sequences y_1^n and y_2^n. Receiver 2 determines the $\hat{\hat{W}}_2$ such that $(u^n(\hat{\hat{W}}_2), y_2^n) \in A_\epsilon^{(n)}$. If there are none or more than one such, an error is declared. Receiver 1 looks for the $(\hat{W}_1, \hat{W}_2)$ such that $(u^n(\hat{W}_2), x^n(\hat{W}_1, \hat{W}_2), y_1^n) \in A_\epsilon^{(n)}$. If there are none or more than one such, an error is declared.

- The remainder of the proof involves showing that the probability of error goes to 0 at both the receivers as long as the rates satisfy (4), and can be looked up in [2].

Outline of converse:

The proof of the converse is provided in [2]. Note that $(W_1, W_2) \leftrightarrow X^n(W_1, W_2) \leftrightarrow Y_1^n \leftrightarrow Y_2^n$. The key idea in the converse is to define the auxiliary random variable U as a function of outputs up to the present time. The basic steps in the proof use Fano's inequality and the fact that conditioning does not increase entropy. $\square$

Remarks:

- By superimposing high-rate information on low-rate information, we can transmit at superior rates compared to time-sharing or maximin procedures [1].

- The bound on the cardinality of U is discussed in [3].

- The proof can use a "subtract-off" or conditioning idea for the better receiver Y_1.

- The cloud center is not explicitly sent.

- For a general BC, the set of rates $\left[R_1 < I(X; Y_1|U),\ R_2 < \min\{I(U; Y_1), I(U; Y_2)\}\right]$ represents an achievable region.

- For BCs, the capacity region depends only on the conditional marginal distributions. To see this, define events $E_1 = \{g_1(Y_1^n) \neq W_1\}$, $E_2 = \{g_2(Y_2^n) \neq W_2\}$ and $E = \{g_1(Y_1^n) \neq W_1 \text{ or } g_2(Y_2^n) \neq W_2\}$. Hence, $\Pr(E_1) \leq \Pr(E)$ and $\Pr(E_1) \leq \Pr(E)$. But $\Pr(E) \leq \Pr(E_1) + \Pr(E_2)$ by the union bound. Thus, $\Pr(E) \to 0 \Leftrightarrow \Pr(E_1) \to 0$ and $\Pr(E_2) \to 0$. Hence, capacity regions found for a particular channel are the same for a larger class of channels having the same conditional marginal distributions.

3.2 The Degraded Gaussian Broadcast Channel

As an extension of the above theorem, consider the degraded Gaussian BC with

$$\begin{aligned} Y_1 &= X + Z_1 \\ Y_2 &= X + Z_2 = Y_1 + Z_2', \end{aligned} \tag{5}$$

where $Z_1 \sim \mathcal{N}(0, N_1)$ and $Z_2' \sim \mathcal{N}(0, N_2 - N_1)$. Here, we have assumed $N_1 < N_2$. Let P be the average input power constraint.

The capacity region of this channel is given by

$$\begin{aligned} R_1 &< C\left(\frac{\alpha P}{N_1}\right) \\ R_2 &< C\left(\frac{(1-\alpha)P}{\alpha P + N_2}\right), \end{aligned} \tag{6}$$

where α may be arbitrarily chosen in $[0, 1]$ and $C(x) := \frac{1}{2}\log(1 + x)$.

Remarks:

- The worse decoder treats the satellite codeword as noise.

- The better decoder can decode Y_2's codeword since it operates at a lower noise level.
- The key ideas in the converse are identification of the superposition variable and the entropy power inequality [4].
- All scalar Gaussian BCs are equivalent to this degraded type, and hence the capacity region for the Gaussian BC is as defined above.

4 Capacity Region Bounds for General Broadcast Channels

The capacity region for a general BC has still not been established except for some special cases [5]; however, several inner and outer bounds to the capacity region have been studied. In this section, we discuss the tightest two such bounds in existing literature - first, the inner bound established by Marton [5] in 1979 and later, the outer bound by Nair and El Gamal [6] published recently at ISIT 2006.

4.1 Marton's Inner Bound

Theorem 4.1. *Let*

$$\begin{aligned} \mathcal{R}_0 &= \{(R_1, R_2) : R_1, R_2 \geq 0, \\ R_1 &< I(U; Y_1), \\ R_2 &< I(V; Y_2), \\ R_1 + R_2 &< I(U; Y_1) + I(V; Y_2) - I(U; V)\} \end{aligned} \tag{7}$$

for some $p(u, v, x)$ on $\mathcal{U} \times \mathcal{V} \times \mathcal{X}$. Then any rate pair $(R_1, R_2) \in \mathcal{R}_0$ is achievable for the discrete memoryless BC $(\mathcal{X}, p(y_1, y_2|x), \mathcal{Y}_1 \times \mathcal{Y}_2)$

Proof. Marton has provided a rather cumbersome proof [5, Theorem 2], which is simplified in [7, 8].
Outline of achievability:

- Fix $p(u, v), p(x|u, v)$. The channel $p(y_1, y_2|x)$ is given, and the idea is to send the auxillary variables u to y_1 and v to y_2.
- *Random Coding:* Generate $2^{nI(U;Y_1)}$ typical u's $\sim p(u)$. Generate $2^{nI(V;Y_2)}$ typical v's $\sim p(v)$. Randomly throw the u's into $\lceil 2^{nR_1} \rceil$ bins and v's into $\lceil 2^{nR_2} \rceil$ bins. For each product bin, find a jointly typical (u^n, v^n) pair. Generate $x^n(u^n, v^n)$ according to $\prod_{k=1}^{n} p(x_k|u_k, v_k)$.
- *Encoding:* To send i to Y_1 and j to y_2, send $x^n(u^n, v^n)$, which is generated by (u^n, v^n) in bin (i, j).
- *Decoding:* Receiver Y_1 finds the u^n that is jointly typical with y_1^n and receiver Y_2 finds v_n typical with y_2^n.

□

Remarks:

- This achievability region is the capacity region if the channel has one deterministic component [5, Theorem 4].

4.2 Deterministic Broadcast Channels

A direct application of Marton's theorem is to compute the achievable region of a deterministic BC. It is shown in [9] that this is the capacity region as well.

Theorem 4.2. *The capacity region of the deterministic memoryless BC with $y_1 = f_1(x)$ and $y_2 = f_2(x)$, is given by the closure of the rate pairs (R_1, R_2) satisfying*

$$\begin{aligned} R_1 &< H(Y_1) \\ R_2 &< H(Y_2) \\ R_1 + R_2 &< H(Y_1, Y_2) \end{aligned} \tag{8}$$

Remarks:

- This can be obtained from the previous theorem by taking $U = Y_1$ and $V = Y_2$.
- The capacity region has a complementary relationship with the Slepian Wolf data compresssion region (Figure 3.)

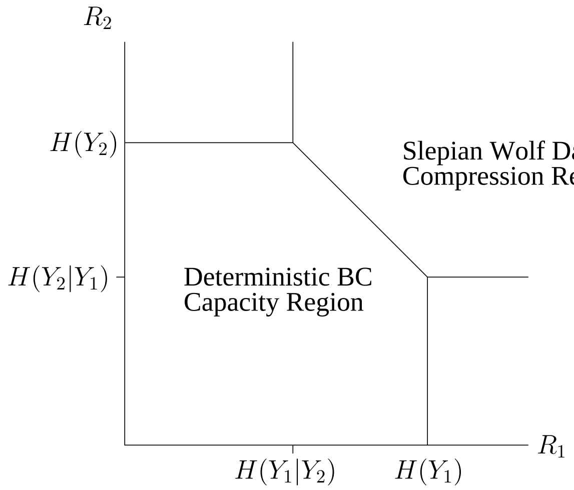

Figure 3: The capacity region for the deterministic BC and the Slepian Wolf data compression region are complementary.

4.3 Nair and Gamal's Outer Bound

This improves upon the earlier bound by Körner and Marton [5, Theorem 5].

Theorem 4.3. *The set of rate pairs (R_1, R_2) satisfying*

$$\begin{aligned} R_1 &< I(U; Y_1) \\ R_2 &< I(V, Y_2) \\ R_1 + R_2 &< \min\{I(U; Y_1) + I(V; Y_2|U), I(V; Y_2) + I(U; Y_1|V)\} \end{aligned} \tag{9}$$

for some choice of joint distributions $p(u, v, x) = p(u, v)p(x|u, v)$ constitutes an outer bound to the capacity region for the discrete memoryless BC. Also, the auxiliary random variables U and V have cardinalitites bounded by $|\mathcal{U}| \leq |\mathcal{X}| + 2$ and $|\mathcal{V}| \leq |\mathcal{X}| + 2$.

Remarks

- The converse proof is given in [6].
- The capacity region for the deterministic BC is obtained by simply taking $U = Y_1$ and $V = Y_2$.

5 Gaussian Vector Broadcast Channels

When a BC is extended to a multiple-input, multiple-output (MIMO)[2] system (having cooperating transmitters), unfortunately, it loses its degradedness in most cases and we see that superposition coding or successive decoding is no longer capacity-achieving. Not to our surprise, the capacity of a general MIMO BC channel remains unsolved. However, for the Gaussian vector BC, the sum capacity region has been characterized [10, 11] and recently, Weingarten et al. [12] have established the entire capacity region.

The general form of a discrete time Gaussian MIMO system is given by $y_i = Hx_i + z_i, \imath = 1, \dots, n$ where $x_i \in \mathbb{C}^t$ is the transmitted vector at time i, $z_i, y_i \in \mathbb{C}^r$ are the noise and corresponding output vectors respectively and $H \in \mathbb{C}^{r \times t}$ is the channel matrix where $h_{k,l}$ denotes the complex channel gain from transmitter l to receiver k. The noise vector z_i has i.i.d values with $z_{k,i} = \mathcal{N}(0,1), k = 1, \dots, r$ and $i = 1 \dots, n$. H is deterministic, fixed and known to both transmittes and receivers. Such a channel in which the transmitters can cooperate and receivers are constrained to decode independently forms a Gaussian MIMO BC.

Caire and Shamai [10] characterized the sum capacity of a BC with two receivers, each equipped with a single antenna, by applying the dirty paper result [13] at the transmitter. This was extended to multiple transmitters and receivers by Yu and Cioffi [11].

5.1 The Dirty Paper region

The capacity of a single-user memoryless channel $P_{Y|X,S}$ with input X, output Y, and interference S, where the interference sequence is noncausally known to the transmitter and unknown to the receiver, is given in terms of the auxiliary random variable U by

$$\sup_{P_{X,U,S}} \{I(U;Y) - I(U;S)\} \tag{10}$$

where the supremum is over all $P_{U,X,S}(x|u,s) = 1\{x = f(u,s)\} P_{U|S}(u|s) P_S(s)$, where $P_S(s)$ is given and $f(u,s)$ is some deterministic function.

In the Gaussian case where $Y = X + S + Z$, the interference S and the noise Z being Gaussian, and X the input under some power constraint $\mathbb{E}[X^2] \leq P$, [13] says that the capacity of the channel is the same as if interference was not present if Z and S are independent.

This can be applied for the MIMO Gaussian BC when choosing codewords for different receivers. The source can order the users and encode each user by treating the previous users as noncausally known interference. This idea of *successive encoding* results in what is known as the *dirty paper region*, which is an achievable region for the Gaussian MIMO BC. The optimality of the dirty paper region has been shown in [10, 11] only for the sum rate. Weingarten et al. have recently shown [12] that dirty paper region achieves capacity for the Gaussian MIMO BC. They also show that superposition coding is optimal for the degraded vector BC while dirty paper coding is optimal for the nondegraded case.

Remarks:

- The approach to compute the dirty paper region is in general complicated and Vishwanath et al. [14] suggest using the duality property of MAC and BC as a simpler procedure to obtain the achievability region.
- Yu and Cioffi [11] show that the sum capacity is a saddle point of a Gaussian mutual information game where a signal player chooses a trasmit covariance matrix to maximize the mutual information and a fictitious noise player chooses a noise correlation to minimize the mutual information.

[2] We use 'MIMO' and 'vector' interchangeably.

References

[1] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 18, no. 1, pp. 2-14, Jan. 1972.

[2] T. M. Cover and J. A. Thomas, "*Elements of Inform. Theory,*" Wiley Interscience, 1991.

[3] R. G. Gallager, "Capacity and coding for degraded broadcast channels," *Problemy Peredachi Informatsii*, vol. 10, no. 3, pp. 3-14, July-Sept. 1974.

[4] P. P. Bergmans, "A simple converse for broadcast channels with additive white Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 20, pp. 279-280, Mar. 1974.

[5] K. Marton, "A coding theorem for the discrete memoryless broadcast channnel," *IEEE Trans. Inform. Theory*, vol. 25, no. 3, pp. 306-311, May 1979.

[6] C. Nair and A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *ISIT* 2006.

[7] T. M. Cover, "Comments on broadcast channels," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2524-2530, Oct. 1998.

[8] A. El Gamal and E. Van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory.* vol. 27, no. 1, pp. 120-122, Jan. 1981.

[9] M. Pinsker, "The capacity region of noiseless broadcast channels," *Problemy Peredachi Informatsii*, vol. 15, pp. 106-108, 1977.

[10] G. Caire and S. Shamai, "On achievable rates in a multi-antenna broadcast downlink," *Proc. 38th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Oct. 2000.

[11] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1875-1892, Sep. 2004.

[12] H. Weingarten, Y. Steinberg and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inform. Theory*, vol. 52, no.9, pp. 3936-3964, Sep. 2006.

[13] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439-441, May 1983.

[14] S. Vishwanath, N. Jindal and A. Goldsmith, "On the capacity of multiple input multiple output broadcast channels," *IEEE Conf. on Communications*, ICC'02, vol. 3, pp. 1444-1450, 2002.